Machine Learning-Based Predictive Modeling for Microbial Inactivation in Nonthermal Processing Technologies

Francesco Donsì [1, 2]*

[1] Department of Industrial Engineering, University of Salerno, Fisciano (SA), Italy

[2] ProdAl Scarl, University of Salerno, Fisciano (SA), Italy

* Corresponding author: fdonsi@unisa.it

Abstract — The effectiveness of microbial inactivation by nonthermal technologies, such as pulsed electric fields (PEF), high-intensity light pulses (HILP), and high-pressure homogenization (HPH), is influenced by multiple factors, including equipment type, operating conditions, process fluid properties, and microorganism resistance. However, the absence of reliable predictive tools for process performance, particularly during industrial scale-up, remains a critical limitation. Unlike conventional models that primarily account for operating conditions, machine learning-based models can integrate additional factors, such as hydrodynamics, liquid media properties, and microorganism-specific characteristics, which are variables that are typically omitted due to their complexity and variability across studies. For instance, in case of HPH, conventional approaches often overlook the influence of homogenization valve geometry or fluid viscosity variations, despite their significant impact on microbial inactivation outcomes. To address these limitations, an extensive dataset of microbial inactivation data from scientific literature was compiled, exceeding 1000 data points for each technology, covering diverse equipment types, process parameters, and microbial species. Data preprocessing included the introduction of dimensionless numbers (e.g., Reynolds, Weber, Capillary, and Cavitation numbers) and covariance analysis to minimize dependence on non-fundamental variables. Various predictive models were tested, including a multilinear regression model as a baseline, two black-box machine learning models, such as artificial neural networks (ANN) and random forests (RF), and a grey-box hybrid model that combines an empirical Weibull model with machine learning techniques. The results demonstrate that machine learning approaches significantly improve predictive accuracy compared to conventional models. While RF models provide faster runtimes without sacrificing performance, the hybrid model further enhances prediction accuracy. Despite these advancements, challenges remain, including the need for additional data and the inclusion of more relevant variables. The predictive tools developed in this study offer a robust framework for optimizing microbial inactivation processes and facilitating the industrial application of PEF, HILP, and HPH.